

Celowość stosowania analizy skupień do oceny zagadnień związanych z jakością energii elektrycznej w sieciach kopalnianych

Streszczenie: Artykuł prezentuje możliwość zastosowania analizy skupień w analizie danych jakości energii elektrycznej. Pierwsza część opisuje ogólnie metody aglomeracji danych oraz wskazuje miejsce analizy skupień w eksploracji danych. Druga część wykazuje cechy analizy skupień sprzyjające zastosowaniu w analizie danych jakości energii elektrycznej. Część trzecia to analiza przypadku zastosowania analizy skupień do wyłonienia zdarzeń napięciowych na podstawie posiadanych danych z stacji elektroenergetycznej wybranego zakładu górniczego zlokalizowanego w Zagłębiu Miedziowym. W artykule wyłonienie zdarzeń napięciowych polega na przeprowadzeniu analizy skupień z zastosowaniem algorytmu K-średnich, czego dokonuje się, dobierając odległości Czebyszewa przy zadanej liczbie skupień.

Słowa kluczowe: jakość energii elektrycznej (JEE), analiza skupień, długoterminowe dane jakości energii elektrycznej, zdarzenia napięciowe, zasilanie zakładów górniczych, generacja rozproszona, odnawialne źródła energii.

The purpose of using the cluster analysis to evaluate the issues related to the quality of the electric power in the mining networks

Summary: The article states that cluster analysis may be used to power quality analysis. First part presents generally the data mining and indicate the placement of cluster analysis in data mining techniques. Second part shows the property indicates of cluster analysis, which indicate the possibility of using it in power quality data analysis. Third part is a case study how to indicates the voltage events by clustering based on power quality data of mining plant in Zagłębie Miedziowe. The indication is made by using K-mean algorithm – Chebyshev distance with the selected number of cluster.

Keywords: power quality, cluster analysis, long term power quality data, voltage events, mining plant supply, distributed generation, renewable sources of energy.

1. Wprowadzenie

Jednym z aktualnych zagadnień związanych z analizą zbiorów danych jest poszukiwanie skutecznych metod ich eksploracji (ang. *data mining*). Przez eksplorację danych należy rozumieć zastosowanie różnych technik analizy obserwowanych zbiorów danych, prowadzące do wyłonienia związków pomiędzy nimi samymi oraz uwarunkowaniami zewnętrznymi. Daje to możliwość poprawy skuteczności interpretacji danych i wsparcia decyzji dotyczących realizacji obserwowanego procesu [2]. Istnieje wiele metod i podziałów eksploracji danych. Jedną z nich jest zaproponowany w [1] podział technik eksploracji danych, zgodnie z którym techniki są następujące:

- klasyfikacja (ang. *classification*),
- regresja (ang. *regression*),
- analiza skupień (ang. *clustering*),
- reguły asocjacyjne (ang. *assosiation rules*).

Niniejszy artykuł dotyczy zastosowania analizy skupień. Ogólnie rzecz ujmując, analiza skupień polega na podziale danych na grupy, w wyniku którego elementy w jednej grupie (skupieniu) są do siebie maksymalnie podobne i z drugiej strony są maksymalnie różne od danych należących do innych grup (skupień) [3]. Analiza skupień ma zatem prowadzić do wyłonienia grup maksymalnie homogenicznych wewnątrz i maksymalnie heterogenicznych w odniesieniu do innych grup. Istnieje wiele metod przeprowadzenia analizy skupień, do najważniejszych należą metody hierarchiczne i niehierarchiczne [6]. Metody hierarchiczne realizowane są poprzez tworzenie hierarchii klasyfikacji – dla n obserwacji tworzy się hierarchię składającą się z n klasyfikacji składających się z 1, 2, 3, ..., n klas. Klasyfikacja zawierająca jedną klasę stanowi zbiór wszystkich obserwacji. Metody niehierarchiczne, zwane również kombinatorycznymi, oparte są na przyporządkowaniu n obiektów do zadanej wcześniej K liczby skupień. Działanie to jest niezależne dla każdej wartości K , ponieważ nie bazuje na wcześniej wyznaczonych skupieniach.

Przeprowadzenie analizy skupień pozwala więc na przejście z dużej liczby danych na rzecz kategorii, co w znacznym stopniu może uprościć analizę, gdyż wnioskowanie o cechach obserwowanego procesu na podstawie dużej liczby danych jest trudne, a niekiedy niemożliwe. Na kategorię (wybrane skupienie) będą składać się poszczególne dane podobne do siebie i różne od danych zawartych w innych skupieniach [6].

2. Dane jakości energii elektrycznej

Analiza jakości energii elektrycznej (JEE) na przestrzeni lat stała się coraz bardziej popularna. Uzasadniona jest już potrzeba prowadzenia tej analizy również w zakładach górniczych Zagłębia Miedziowego ze względu na korzyści z niej płynące [3, 4]. Wykorzystywany jest nie jest tylko monitoring sieci elektroenergetycznych w trakcie audytów

czy tygodniowych analiz JEE, ale duże zakłady przemysłowe sięgają po stałe rozwiązania. Coraz popularniejsze stają się stacjonarne systemy monitorowania jakości energii elektrycznej w ujęciu globalnym przedsiębiorstwa lub lokalnie dla wybranych newralgicznych maszyn i urządzeń elektrycznych. Ciągły monitoring parametrów jakościowych prowadzi do powstawania bardzo dużej ilości danych. Analizy parametrów wymaganych przez właściwe rozporządzenie [9] czy normę [7] sprowadza się do analizy następujących parametrów zaburzeń napięcia: częstotliwości zasilania (f), wartości skutecznej napięcia (U), migotania światła (P_{it}), asymetrii napięcia (k_{U2}) czy zawartości harmonicznych (THD_U , U_{h2} ... U_{h50}). Dane te przed analizą są agregowane z odpowiednim interwałem czasowym określonym w [8] (10 minut dla wszystkich wyżej wskazanych z wyjątkiem $P_{it} - 2h$), co daje w przeciągu tylko jednego tygodnia analizy 1008 danych 10-minutowych dla każdego z parametrów i 84 dane 2-godzinne. A zatem analiza danych z dłuższego okresu czasu (długoterminowe dane jakości energii elektrycznej) staje się trudnym zadaniem ze względu na ilość danych. Korzystne wydaje się więc zastosowanie analizy skupień do utworzenia grup danych jakości energii elektrycznej o podobnych cechach.

3. Opis obiektu badań i przeprowadzonych pomiarów jakości energii elektrycznej

Badania wykonano w polach liniowych 110 kV stacji zasilającej 110 kV/6 kV należącej do wybranego Oddziału Zakładów Górniczych Zagłębia Miedziowego. Dodatkowo na poziomie średniego napięcia przyłączona jest lokalna generacja – bloki parowo-gazowe. Zasilanie odbiorników jest przeprowadzone w układzie promieniowym. Sieci zasilające 110 kV pracują z bezpośrednio uziemionym punktem neutralnym.

Na potrzeby analizy pod uwagę wzięto parametry jakości energii elektrycznej w trzech liniach zasilających 110 kV (oznaczonych jako S1, S2, S3):

- S-1 jest typową linią, do której przyłączone są tylko odbiory (ciągły pobór mocy czynnej),
- S-2 jest linią, do której przyłączona jest znaczna liczba odbiorników oraz generacja równoważąca, a niekiedy przewyższająca zapotrzebowanie (blok 1),
- S-3 jest linią, która przez większość czasu charakteryzuje się większą produkcją energii pochodzącej z lokalnej generacji (blok 2) niż obciążeniem.

Rejestrację parametrów wykonano w okresie jesienno-zimowym od 21 listopada do 18 grudnia. Do rejestracji parametrów jakościowych wykorzystano trzy analizatory jakości energii klasy A. Pomiar wykonywany był w układzie pośrednim z wykorzystaniem stacyjnych przekładników prądowych i napięciowych. Metodykę pomiarów i czasy agregacji wykonano zgodnie z klasą A normy [8]. Dokonano pomiarów synchronicznych z wykorzystaniem modułu GPS w trzech liniach równolegle. Analizowano wszystkie parametry jakości energii elektrycznej w każdej linii (S-1, S-2, S-3).

4. Zastosowanie algorytmu K-średnich w analizie danych jakości energii elektrycznej

Przeprowadzenie oceny zgodności parametrów jakości energii elektrycznej z właściwym rozporządzeniem [9] lub normą [7] powinno być poprzedzone wyeliminowaniem danych zagregowanych zawierających w sobie zdarzenia napięciowe np. skojarzone ze stanem awaryjnym (tzw. system flagowania). Zadanie to można wykonać manualnie na podstawie posiadanej wiedzy. Autorzy zaobserwowali możliwość zastosowania niehierarchicznej analizy skupień przy pomocy algorytmu K-średnich, wybierając miarę odległości Czebyszewa do wyodrębnienia zdarzeń napięciowych jako grup danych znacznie odbiegające od pozostałych.

Analizę skupień przeprowadzono z wykorzystaniem programu Statistica. Dokonano uogólnionej analizy skupień za pomocą algorytmu K-średnich. Niezbędne było określenie odległości, względem której algorytm przydziela dane do skupienia. Możliwe jest zastosowanie: odległości euklidesowej, kwadratu odległości euklidesowej, odległości miejskiej (Manhattan, City block) oraz odległości Czebyszewa. Odległości te określone są przez następujące zależności [10]:

- odległość euklidesowa: $odległość(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$,
- kwadrat odległości euklidesowej: $odległość(x, y) = \sum_i (x_i - y_i)^2$,
- odległości miejska (Manhattan): $odległość(x, y) = \sum_i |x_i - y_i|$,
- odległość Czebyszewa: $odległość(x, y) = \text{Maksimum } |x_i - y_i|$,

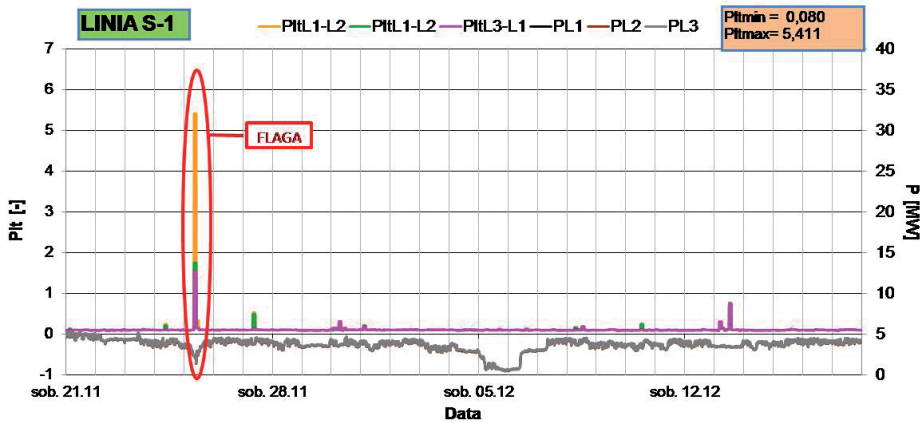
Do wyłonienia zdarzeń napięciowych wybrano odległość Czebyszewa. Przeprowadzono trzy analizy skupień (dla każdej linii osobno) w oparciu o algorytm K-średnich z wykorzystaniem odległości Czebyszewa o zadanej wejściowej liczbie skupień 2. Dane były standaryzowane, tzn. przybierały wartości z przedziału [0, 1]. Przydział do odpowiednich skupień określony został na podstawie maksymalizacji odległości skupień. Do analizy brane były pod uwagę parametry: częstotliwość zasilania (f), wartości skuteczne napięcia (U), współczynnik migotania światła (P_{lt}), współczynnik asymetrii napięcia (k_{u2}), zawartość harmonicznych w napięciu (THD_U).

5. Wyniki pomiarów i przeprowadzonej analizy skupień

W celu sprawdzenia możliwości stosowania analizy skupień jako narzędzia do wyłonienia zdarzeń napięciowych z długoterminowych danych jakości energii elektrycznej porównano analizy wykonane na podstawie wiedzy eksperckiej, tj. wyników analizy danych wykonanych przez autorów podczas analizy zgodności parametrów z właściwym

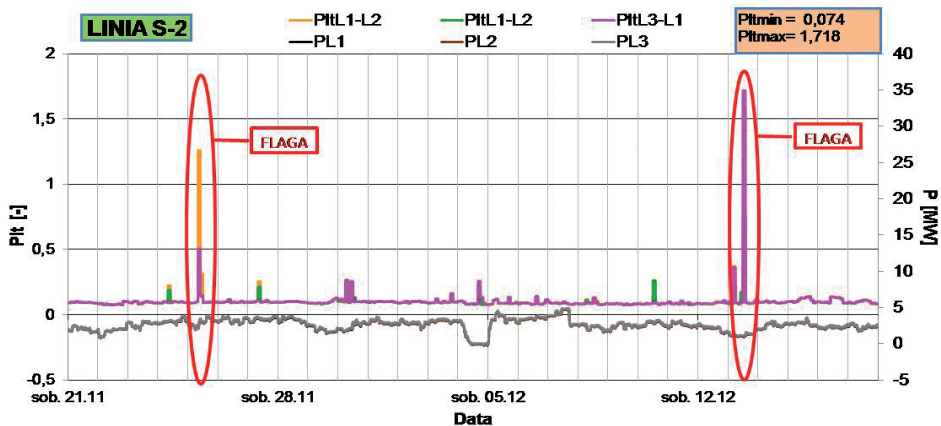
rozporządzeniem [9] i normą [7], oraz przeprowadzonej analizy skupień na podstawie wydzielonej klasyfikacji danych, tj. przynależności do wyłonionego skupienia.

Na rysunkach 1–3 wskazano zmienność współczynnika wahań napięcia P_{it} w czasie pomiarów, odpowiednio w liniach S-1, S-2, S-3 (pionowa oś lewa rysunków). Dodatkowo każdy z rysunków prezentuje zmienność obciążenia w danej linii (pionowa oś prawa rysunków). Na podstawie wiedzy eksperckiej oflagowano dwa zdarzenia napięciowe związane ze stanami przejściowymi w sieci elektroenergetycznej. W konsekwencji w ocenie jakości energii elektrycznej nie uwzględniano danych skojarzonych z wyłonionymi zdarzeniami.



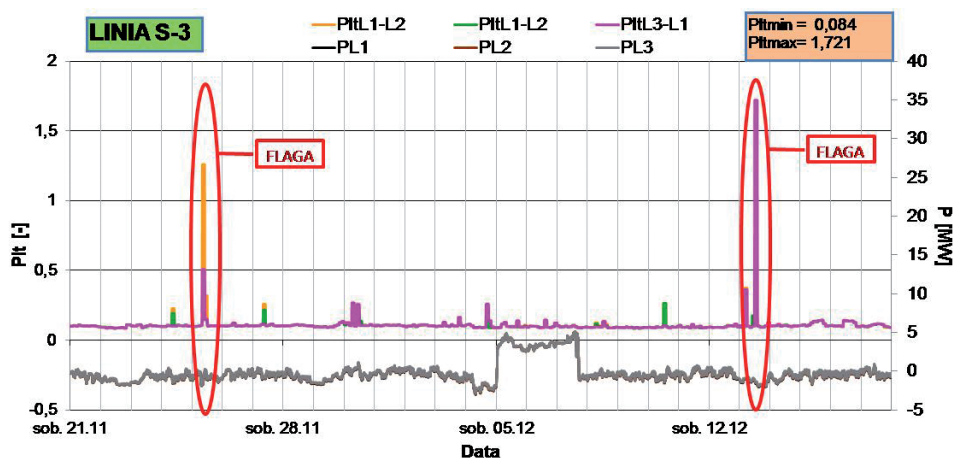
Rys. 1. Analiza wahań napięcia i zmienności obciążenia w linii S-1

Źródło: oprac. własne.



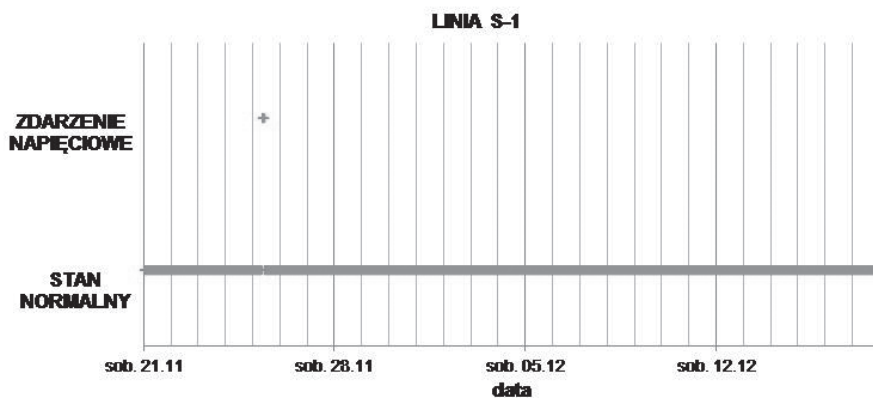
Rys. 2. Analiza wahań napięcia i zmienności obciążenia w linii S-2

Źródło: oprac. własne.



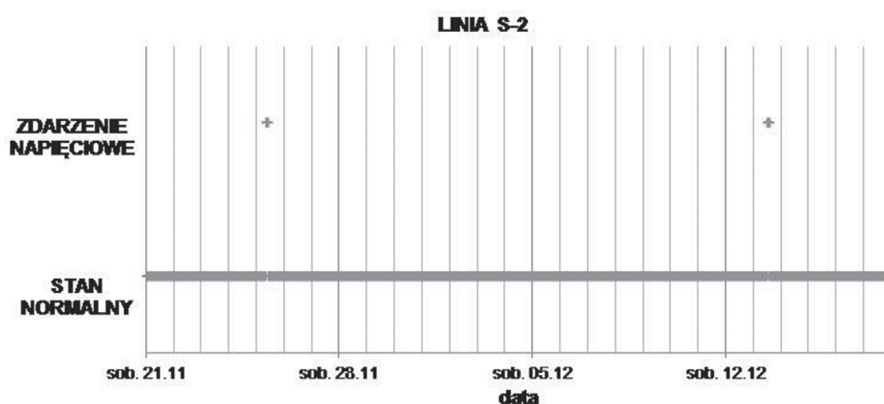
Źródło: oprac. własne.

Na rysunkach 4-6 wskazano wyniki przeprowadzonej analizy skupień za pomocą algorytmu K-średnich zgodnie z podanymi wcześniej założeniami odpowiednio dla linii S-1, S-2, S-3. Dane zostały przypisane do dwóch skupień: skupienie 1 – stan normalny, skupienie 2 – zdarzenie napięciowe.



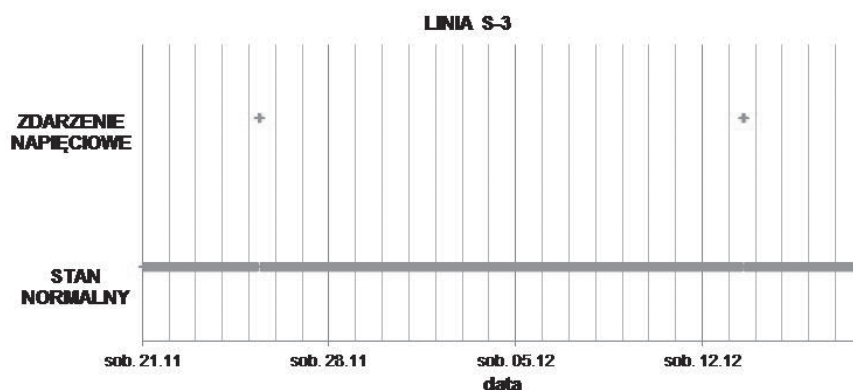
Rys. 4. Wyniki analizy skupień w linii S-1

Źródło: oprac. własne.



Rys. 5. Wyniki analizy skupień w linii S-2

Źródło: oprac. własne.



Rys. 6. Wyniki analizy skupień w linii S-3

Źródło: oprac. własne.

Należy zwrócić uwagę, że te same dane zostały oflagowane na podstawie wiedzy eksperckiej (rysunki 1–3) oraz wyodrębnione na podstawie analizy skupień (rysunki 4–6). Potwierdza to możliwość stosowania analizy skupień z zastosowaniem algorytmu K-średnich i doбором odległości Czebyszewa jako narzędzia do wstępnej analizy danych i wydzielenia zdarzeń napięciowych z dużej ilości danych pomiarowych bez konieczności posiadania wiedzy eksperta. Autorzy przeprowadzili podobne analizy dla innych parametrów jakości energii elektrycznej, potwierdziły one zasadność stosowania analizy skupień do wyłonienia zdarzeń napięciowych.

6. Podsumowanie

Wyniki przedstawione w artykule potwierdzają możliwość zastosowania analizy skupień do wyłonienia danych należących do zdarzeń napięciowych, podlegających zasadzie flagowania, jako pierwszego etapu analizy danych jakości energii elektrycznej. W dalszym zastosowaniu analizy skupień w analizie danych JEE (nie tylko w celu wyłonienia zdarzeń napięciowych) upatruje się następujące korzyści:

- redukcję dużej ilości danych na rzecz kilku podstawowych kategorii, które można uznać za przedmiot dalszych analiz,
- uzyskanie jednorodnych grup, dla których można określić zasadnicze cechy,
- uzyskanie klasyfikacji stanów pracy danego obiektu.

W dalszych badaniach planuje się wykorzystać analizę skupień do porównania różnych stanów pracy, w jakich znajduje się badany obiekt, np. wskazania wpływu generacji rozproszonej na warunki pracy elektroenergetycznych sieci kopalnianych.

Literatura

- [1] CIGRE Brochure 292, *Data mining techniques and applications in the power transmission field*, CIGRE, 2006.
- [2] Larose D., *Metody i modele eksploracji danych*, Wydawnictwo PWN, Warszawa 2005.
- [3] Jasiński M., Jasiński M., Jasiński Ł., *Potrzeba analizy jakości energii elektrycznej w kopalniach rud miedzi. Generacja – Przesył – Wykorzystanie*, GPW 2015 [Dokument elektroniczny], praca zbiorowa pod red. Macieja Gwoździewicza, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2015.
- [4] Jasiński M., Kaczorowska D., Jasińska L., Jasiński Ł., *Wpływ jakości energii elektrycznej na funkcjonowanie kopalni*, „Zeszyty Naukowe Uczelni Jana Wyżykowskiego. Studia z Nauk Technicznych” 2016, z. 5.
- [5] Jasiński M., Sikorski T., Karpiński J., Zenger M., *Analizy statystyczne długoterminowych danych jakości energii elektrycznej. Generacja – Przesył – Wykorzystanie*, GPW 2016 [Dokument elektroniczny], praca zbiorowa pod red. Macieja Gwoździewicza, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2016.
- [6] Migut G., *Zastosowanie technik analizy skupień i drzew decyzyjnych do segmentacji rynku*, http://media.statsoft.nazwa.pl/_old_dnn/downloads/zastosowanie_teknik.pdf, stan z dnia 1.05.2017.
- [7] *PN-EN 50160. Parametry napięcia zasilającego w publicznych sieciach rozdzielczych*, 2010.
- [8] *PN-EN 61000 4-30: Kompatybilność elektromagnetyczna (EMC) Część 4-30: Metody badań i pomiarów: Metody pomiaru jakości energii*, 2011.
- [9] Rozporządzenie Ministra Gospodarki z dnia 4 maja 2007 r. w sprawie szczegółowych warunków funkcjonowania systemu elektroenergetycznego, ze zmianami z dnia 21 sierpnia 2008 r. (Dz.U. z 2008 r. nr 162, poz. 1005).
- [10] StatSoft, *Elektroniczny Podręcznik Statystyki PL*, <http://www.statsoft.pl/textbook/stathome.html>, 2016, stan z dnia 14.06.2017.
- [11] Wierchoń S., Kłopotek M., *Algorytmy analizy skupień*, WNT, Warszawa 2015.